

Credit Card Fraud Predictive Modeling and Deployment

**Final Report**

Higher Diploma in Science in Data Analytics

Word Count: 7500

Author: Ciaran Finnegan / 10524150

E-mail: [10524150@mydbs.ie](mailto:10524150@mydbs.ie) / [ciaran@feefinnegan.com](mailto:ciaran@feefinnegan.com)

Supervisor: Dr Shahram Azizi Sazi

Report Submission Date: 25th September 2020

**Abstract**

This is, at maximum, a quarter page summary of the project…

**Acknowledgments**

I wish to acknowledge the..

# Contents

[Contents 2](#_Toc48931159)

[1. Introduction 4](#_Toc48931160)

[1.1. What the Project Aimed to Deliver 4](#_Toc48931161)

[1.2. How the Project Delivery was Implemented 4](#_Toc48931162)

[2. Background / Literature Review 5](#_Toc48931163)

[2.1. Credit Card Fraud Detection: Further Research on Predictive Models 5](#_Toc48931164)

[3. Requirements: Specification and Design 6](#_Toc48931165)

[3.1. High Level Project Requirements 6](#_Toc48931166)

[3.2. Project Architecture Diagram 6](#_Toc48931167)

[3.3. High Level Project Design 7](#_Toc48931168)

[3.3.1. Prototype Development 7](#_Toc48931173)

[3.3.2. Final Project Deliverable 7](#_Toc48931174)

[4. Project Implementation(1) – Azure Modelling 8](#_Toc48931179)

[4.1. The Machine Learning Workflow 8](#_Toc48931180)

[4.2. Credit Card Fraud Dataset – Analysis and Preparation 8](#_Toc48931181)

[4.3. Credit Card Fraud – Building the Azure Model 8](#_Toc48931182)

[4.4. Credit Card Fraud – Deploying the Azure Model 8](#_Toc48931183)

[5. Project Implementation(2) – Shiny R Dashboard UI 9](#_Toc48931184)

[5.1. Data Visualisations in a Shiny Dashboard 9](#_Toc48931185)

[5.2. Credit Card Fraud – UI to Check Fraud Predictions 9](#_Toc48931186)

[5.3. Shiny UI – Hosted Application 9](#_Toc48931187)

[6. Testing and Results 10](#_Toc48931188)

[6.1. User Story ‘Demos’ – Test Results and ‘Feedback’ 10](#_Toc48931189)

[7.1.1. User Story 4: Initial Data Modelling – Review and Evaluation 10](#_Toc48931194)

[7.1.2. User Story 5: Basic Shiny App – Review and Evaluation 10](#_Toc48931195)

[7.1.3. User Story 6: Integrated Prototype – Review and Evaluation 10](#_Toc48931196)

[7.1.4. User Story 7: Enhanced Modelling – Review and Evaluation 10](#_Toc48931197)

[7.1.5. User Story 8: Enhanced UI – Review and Evaluation 10](#_Toc48931198)

[7.1.6. User Story 9: Presentation Preparation – Review and Evaluation 11](#_Toc48931199)

[6.2. Final Project Assessment 12](#_Toc48931200)

[6.3. Project Plan 2020: Final Status – 25th September 2020 13](#_Toc48931201)

[7. Project Location and User Guide 14](#_Toc48931202)

[7.1. Credit Card Fraud Application: Prototype Location 14](#_Toc48931203)

[7.2. Credit Card Fraud Application: User Guide (Final Project) 14](#_Toc48931204)

[8. Project Conclusions 15](#_Toc48931205)

[8.1. Where Project Goals Achieved? 15](#_Toc48931206)

[8.2. Future Design/Deployment Considerations 15](#_Toc48931207)

[9. Appendices 16](#_Toc48931208)

[9.1. Azure Generated Code Segments 16](#_Toc48931209)

[9.2. Shiny R Application Code Files 18](#_Toc48931210)

[13.2.1. Diagram: The RStudio Cloud Environment 18](#_Toc48931219)

[13.2.2. The Shiny UI Code 18](#_Toc48931220)

[13.2.3. The R Code Parsing Data and Invoking UI 18](#_Toc48931221)

[9.3. Azure Machine Learning Classic Studio Experiments 19](#_Toc48931222)

[9.4. Credit Card Fraud Datasets 20](#_Toc48931223)

[10. References / Bibliography - Interim Report 21](#_Toc48931224)

# Introduction

## What the Project Aimed to Deliver

The artefact at the end of this project is an application that invokes a bespoke predictive model and provides a user with an online interface to retrieve a score for whether a given credit card transaction is likely to be fraudulent.

The Interim Report…

<image>

## How the Project Delivery was Implemented

The Interim Report…

The project .

* The entire remaining 100K+ rows of the credit card fraud dataset will be used to generate the final predictive model.

# Background / Literature Review

## Credit Card Fraud Detection: Further Research on Predictive Models

The Interim Report…

<image>

One-pager..

As of mid-August 2020, ..

* The...

# Requirements: Specification and Design

## High Level Project Requirements

To provide an synopsis of the project requirements detailed in the Interim Report..

* A predictive model for Credit Card fraud detection ..
* A ..
* All ..

## Project Architecture Diagram

*Figure: High Level Application Architecture Diagram*

A close up of a map

Description automatically generated

## High Level Project Design

One-page description with a diagram…

The Interim report provided a detailed overview of the User Stories used to map out the design and implementation of this project.



### Prototype Development

*Initial Basic Modelling in Azure ML Studio (Classic)*

The,,

<image>

*Basic UI Deployment*

The..

<image>

### Final Project Deliverable

*Enhanced Modelling in Azure ML Studio (Classic)*

The,,

<image>

*Enhanced UI Deployment*

The..

<image>



# Project Implementation(1) – Azure Modelling

## The Machine Learning Workflow

The ..

## Credit Card Fraud Dataset – Analysis and Preparation

The..

<images>

## Credit Card Fraud – Building the Azure Model

The..

<images>

## Credit Card Fraud – Deploying the Azure Model

The..

<images>

# Project Implementation(2) – Shiny R Dashboard UI

## Data Visualisations in a Shiny Dashboard

The prototype for this project application..

The..

<images>

## Credit Card Fraud – UI to Check Fraud Predictions

The..

<images>

## Shiny UI – Hosted Application

The..

<images>

# Testing and Results

## User Story ‘Demos’ – Test Results and ‘Feedback’

The prototype for this project..



### User Story 4: Initial Data Modelling – Review and Evaluation

Goal:

Assessment of robustness of code and functionality delivered:

### User Story 5: Basic Shiny App – Review and Evaluation

Goal:

Assessment of robustness of code and functionality delivered:

### User Story 6: Integrated Prototype – Review and Evaluation

Goal:

Assessment of robustness of code and functionality delivered:

### User Story 7: Enhanced Modelling – Review and Evaluation

Goal:

Assessment of robustness of code and functionality delivered:

### User Story 8: Enhanced UI – Review and Evaluation

Goal:

Assessment of robustness of code and functionality delivered:

### User Story 9: Presentation Preparation – Review and Evaluation

Goal:

Assessment of robustness of code and functionality delivered:

## Final Project Assessment

The project is intended to demonstrate …

## Project Plan 2020: Final Status – 25th September 2020

**(Produced using the Team Gantt online portal)**

A screenshot of text

Description automatically generated

# Project Location and User Guide

## Credit Card Fraud Application: Prototype Location

The prototype for this project application is currently hosted on *shinyapps.io* and the UI can be accessed through this URL;

<https://ciaran-finnegan.shinyapps.io/DBS_CCFraudRShinyApp_10524150/>

## Credit Card Fraud Application: User Guide (Final Project)

The prototype is intended to demonstrate progress to date and provide tangible evidence of the end goals of this project.

*Figure: Final Project*

A screenshot of a cell phone

Description automatically generated

A User Guide, in Microsoft PowerPoint format, is embedded with this report, and has also been submitted separately…

# Project Conclusions

## Where Project Goals Achieved?

The User Stories ...

## Future Design/Deployment Considerations

The Interim Report.....

● Any changes from the interim report should be discussed and justified.

● The student should reflect on the learning experiences gained in doing the project and its relevance to on--going progress as a learner and future practising IT professional.

● This section should also provide a starting point for another student to continue the work.

# Appendices

## Azure Generated Code Segments

The Azure Machine Learning Studio auto-generates codes segments in C#, Python 3.6, and R to access both Azure hosted datastores and invoke APIs to Azure hosted Rest Endpoints for deployed models.

Below are examples of code snippets which have been incorporated into my project.

This code segments reads a subset of the Credit Card dataset and uses the output to generate data visualisations in the R Shiny App.

*Figure: R code snippet to read Azure hosted datastore*

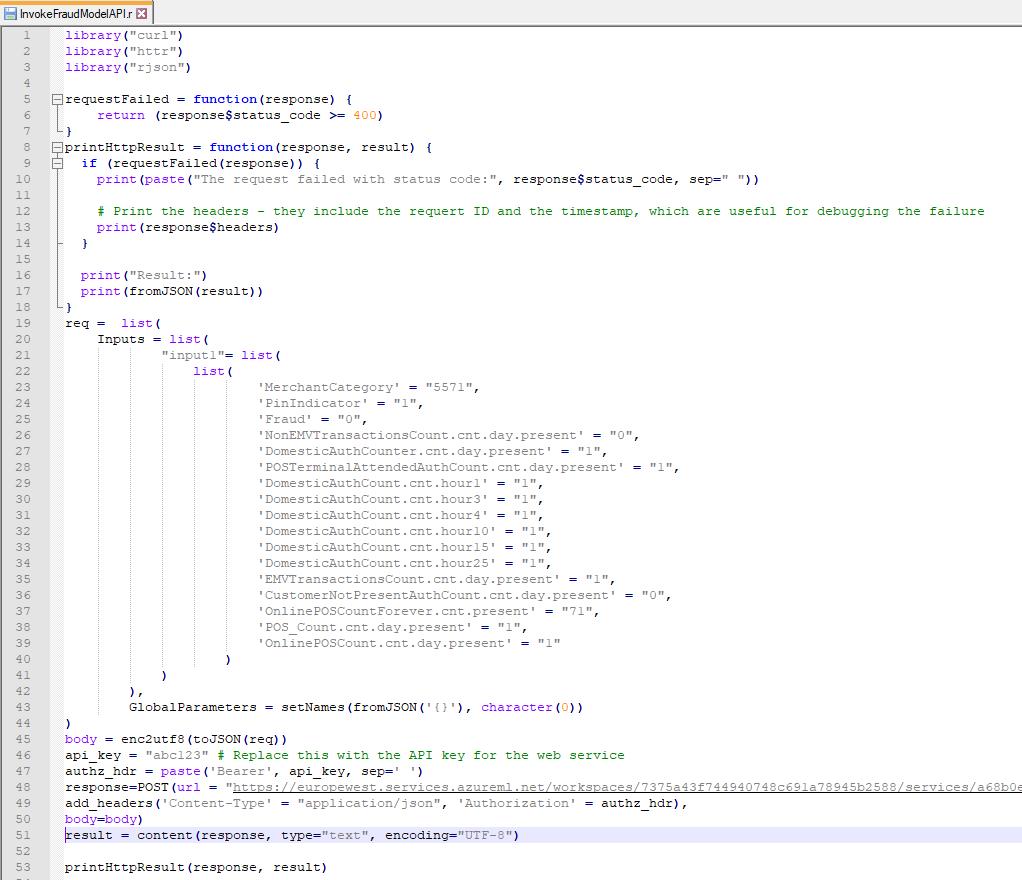
A screenshot of a social media post

Description automatically generated

This code segment invokes the API to pass attributes to the Fraud detection model, hosted in Azure, and returns a prediction score.

(Line 48 has been truncated slightly).

*Figure: R code snippet to read Azure hosted REST Endpoint for Fraud Model*



## Shiny R Application Code Files



### Diagram: The RStudio Cloud Environment

A screenshot of a social media post

Description automatically generated

### The Shiny UI Code

The..

### The R Code Parsing Data and Invoking UI

The..

## Azure Machine Learning Classic Studio Experiments

The final…

## Credit Card Fraud Datasets

The datasets..

# References / Bibliography - Interim Report

1Microsoft Tutorial, 2020. *What Is ML Studio (Classic) - Azure*. [online] Docs.microsoft.com. Available at: <https://docs.microsoft.com/en-us/azure/machine-learning/studio/what-is-ml-studio> [Accessed 19 June 2020].

2Krzemiński, D., 2018. *Create Outstanding R Shiny Dashboards With The Semantic.Dashboard Package - Appsilon Data Science | End­ To­ End Data Science Solutions*. [online] Appsilon Data Science | End­ to­ End Data Science Solutions. Available at: <https://appsilon.com/create-outstanding-dashboards-with-the-new-semantic-dashboard-package/> [Accessed 28 June 2020].

3Parsons, T., 2019. *When To Use Waterfall Vs. Agile | Macadamian*. [online] Macadamian. Available at: <https://www.macadamian.com/learn/when-to-use-waterfall-vs-agile/> [Accessed 10 July 2020].

4Stanton, K., 2019. *Card Fraud In Europe - It's Still Increasing..*. [online] Finextra Research. Available at: <https://www.finextra.com/blogposting/16824/card-fraud-in-europe---its-still-increasing> [Accessed 4 August 2020].

5Cherowbrier, J., 2019. *Forecast Of Card Payments Per Day 2006, 2016 And 2026 Kingdom | Statista*. [online] Statista. Available at: <https://www.statista.com/statistics/719708/card-payments-per-day-forecast-united-kingdom/> [Accessed 7 August 2020].

6ICA, n.d. *What Is Customer Due Diligence (CDD)? | ICA*. [online] Int-comp.org. Available at: <https://www.int-comp.org/careers/your-career-in-aml/what-is-customer-due-diligence-cdd/> [Accessed 7 August 2020].

7Machine Learning Group - ULB, 2018. *Credit Card Fraud Detection*. [online] Kaggle.com. Available at: <https://www.kaggle.com/mlg-ulb/creditcardfraud> [Accessed 13 August 2020].

8Preda, G., 2018. *Credit Card Fraud Detection Predictive Models*. [online] Kaggle.com. Available at: <https://www.kaggle.com/gpreda/credit-card-fraud-detection-predictive-models#Data-exploration> [Accessed 13 June 2020].

9Charminda D, 2020. *Credit Card Fraud Detection: A Detailed Study*. [online] Kaggle.com. Available at: <https://www.kaggle.com/chandrimad31/credit-card-fraud-detection-a-detailed-study> [Accessed 13 July 2020].

10Escapa, C., 2018. *AWS Analytics Services Explained: From Data Lakes To Machine Learning | Amazon Web Services*. [online] Amazon Web Services. Available at: <https://aws.amazon.com/blogs/apn/aws-analytics-services-explained-from-data-lakes-to-machine-learning/> [Accessed 7 August 2020].

11Sharma, A., 2020. *How To Use Google Colab For Deep Learning And Machine Learning*. [online] Analytics Vidhya. Available at: <https://www.analyticsvidhya.com/blog/2020/03/google-colab-machine-learning-deep-learning/#1> [Accessed 7 August 2020].

12Microsoft Tutorial, 2020. *ML Studio (Classic): Quickstart: Create A Data Science Experiment - Azure*. [online] Docs.microsoft.com. Available at: <https://docs.microsoft.com/en-us/azure/machine-learning/studio/create-experiment> [Accessed 7 August 2020].

13Kipp, A., 2017. *Shiny - Shinyapps.Io - Getting Started*. [online] Shiny.rstudio.com. Available at: <https://shiny.rstudio.com/articles/shinyapps.html> [Accessed 7 July 2020].

14Highsmith, J., 2007. *Agile Project Management*. 23rd ed. Boston, Mass.;Munich[u.a.]: Addison-Wesley.

15Beck, et al., 2001. *Manifesto For Agile Software Development*. [online] Agilemanifesto.org. Available at: <https://agilemanifesto.org/> [Accessed 7 August 2018].

16Agile Alliance, n.d. *What Are User Stories?*. [online] Agile Alliance. Available at: <https://www.agilealliance.org/glossary/user-stories> [Accessed 7 March 2018].

17Sachdeva, S., 2016. Scrum Methodology. *International Journal Of Engineering And Computer Science*,.

18Huether, D., 2017. *The Definition Of Done*. [online] LeadingAgile. Available at: <https://www.leadingagile.com/2017/02/definition-of-done/#:~:text=The%20definition%20of%20done%20(DoD,%2C%20team%2C%20or%20consuming%20system.> [Accessed 7 March 2018].

19Verner, T., 2019. *Microsoft Azure AI Engineer: Developing ML Pipelines In Microsoft Azure*. [online] Pluralsight.com. Available at: <https://www.pluralsight.com/courses/microsoft-azure-developing-ml-pipelines> [Accessed 7 June 2020].

20Deliwala, S., 2019. *Streamlit 101: An In-Depth Introduction*. [online] Medium. Available at: <https://towardsdatascience.com/streamlit-101-an-in-depth-introduction-fc8aad9492f2> [Accessed 7 August 2020].